REMARKS

In paragraph 3, on page 2 of the Office Action, the Examiner rejects Claims 85-106 under 35 U.S.C. § 112, first paragraph.

Specifically, the Examiner states that the recitation "at least about 10 nucleotides in length" is not supported in the specification, i.e., the Examiner contends that the specification (see page 10, and original Claim 5) only supports "about 10 to about 20 nucleotides".

For the following reasons, Applicants respectfully traverse the Examiner's rejection.

The test for sufficiency of support in an application is whether the disclosure of the application relied upon "reasonably conveys to the artisan that the inventor had possession at that time of the later claimed subject matter" (*In re Kaslow*, 217 USPQ 1089 (Fed. Cir. 1983)).

As stated by the Federal Circuit in *Ralston Purina Company v. Far-Mar-Co., Inc.*, 227 USPQ 177 (Fed. Cir. 1985):

> Far-Mar-Co cites several range cases to support its argument that ranges found in the applicant's claim language must correspond exactly to ranges disclosed in the parent. These cases are not on point. The facts in these cases precluded a determination that one skilled in the art could derive the claim limitations from the parent, due to a number of different factors, e.g., the unpredictable nature of the art, *In re Sichert*, 566 F.2d 1154, 196 USPQ 209 (CCPA 1977); failure to distinguish one process from another, *In re MacLean*, 454 F.2d 756, 172 USPQ 494 (CCPA 1972); the addition of a critical limitation, *In re Blaser*, 556 F.2d 534, 194 USPQ 122 (CCPA 1977); failure to define a critical term, *In re Lukach*, 442 F.2d 967, 169 USPQ 795 (CCPA 1971); and use of a list that did not contain the claimed substance. *In re Ahlbrecht*, 435 F.2d 908, 168 USPQ 293 (CCPA 1971). In addition, a predecessor to

> this court has held "that a claim may be broader than the specific embodiment disclosed in a specification is in itself of no moment." *In re Rasmussen*, 650 F.2d 1212, 1215, 211 USPQ 323, 326 (CCPA 1981)...
>
> [T]he court did not clearly err in determining that the parent's disclosure adequately supported the water ranges of "at least about 25% by weight," and "at least 25% by weight." [where the examples of the parent application showed a water content that equaled 25% and 27%]. (Emphasis added)

There is no unpredictable nature with respect to the size of the oligonucleotide of the present invention, as it is well-known in the art that probes can be greater than 20 nucleotides. Further, the upper limit of the size of the probe is not relevant to distinguishing over the prior art, i.e., such is not a critical limitation or a critical term.

The specification, at page 5, teaches that the nucleic acids may be variable in length, and that in one embodiment, they are from about 10 to about 20 nucleotides in length. Thus, the range of "about 10 to about 20" is merely an example of one embodiment.

Clearly, the examples in the present application (see also, e.g., Claim 88) teach that the oligonucleotides can be 10 or greater nucleotides in length, and can be greater than 20 nucleotides (e.g., positions 11821-11844 of SEQ ID NO:1 (i.e., 24 amino acids) and positions 12945-12924 of SEQ ID NO:1 (i.e., 22 amino acids)).

Similarly, in *In re Wertheim*, 191 USPQ 90 (CCPA 1976) the asserted claims covered a range ("solids level of at least 35%"), whereas the specification disclosed a range ("concentrated . . . until a concentration of 25 to 60% solid matter is reached").. The court held that the claim found

written description in the specification even though the claims was not repeated verbatim in the specification.

Hence, it is clear that the specification clearly conveys to one skilled in the art "at least about 10" as recited in the claims.

In addition, in paragraph 5, on page 4 of the Office Action, the Examiner states that the method of Claim 101 requires one to perform hybridization reactions between a probe of undefined sequence to a target sequence found in any of a variety of "sugar-pathway genes". The Examiner contends that the specification fails to provide an adequate written description of the coding sequence for "sugar-pathway genes" as found in any of the encompassed microorganisms, nor an adequate written description of the probes and primers that would be used in the hybridization and detection aspects of the method.

For the following reasons, Applicants respectfully traverse the Examiner's rejection.

It is well-settled law that an application need not teach, and preferably omits, that which is well-known in the art (*Hybritech Inc. v. Monoclonal Antibodies, Inc.*, 231 USPQ 81 (Fed. Cir. 1986)).

Applicants respectfully submit that the sequences of oligonucleotide molecules which hybridize to the "sugar-pathway genes specific to the bacterial strains to be detected" were well-known to one skilled in the art, and/or ones skilled in the art would have been able to identify oligonucleotide molecules capable of hybridizing thereto without undue experimentation (see the last paragraph at page 6 of the present specification).

Moreover, Applicants submitted with the Amendment After Final filed December 23, 2002, a list of citations dated prior to 1997 (Appendix B) and copies of related references (Appendix C) which disclose a range of genes associated with sugar biosynthetic pathways. These citations show that sugar-pathway genes were well-known in the art, and formed part of the common general knowledge. As a result, one skilled in the art would have been able to use her common general knowledge and routine techniques available in the art to produce oligonucleotide molecules capable of hybridizing to such genes.

In the outstanding Office, the Examiner has apparently overlooked these references in maintaining the outstanding rejection.

Applicants wish to add the following citation to those previously submitted:

> Keenylside et al "A Novel Pathway for O-Polysaccharide Biosynthesis in *Salmonella enterica Serovar Borreze*", *Journal of Biological Chemistry*, Volumn 271, No. 45, pages 28581-28592 (1996); a copy of which is enclosed herewith.

The above reference clearly shows that sugar-pathway genes were well-known in the art, and formed part of the common general knowledge.

Applicants respectfully direct the Examiner's attention to the specific instances of the use of the term "pathway" in "*Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, Vol. 1", particularly at pages 504 and 507. At these sections, the use of the term "pathway" clearly relates to the biosynthetic processes involved in the synthesis of sugars and

more complex carbohydrates, and that by the additional use of the terms "sugar" and "gene", there is explicitly described to the person skilled in the art a series of genes directly associated with these biosynthetic pathways. This is even more clearly described at pages 63 and 64 in "Physiology of the Bacterial Cell, A Molecular Approach" (Neiderhardt et al).

Furthermore, in the article by Keenylside et al, it is stated at page 28581, second column, last paragraph:

> Enzymes encoded by the *rfb*$_{O:54}$ cluster direct the synthesis of an N-acetylmannosamine (ManNAc)-containing O-polysaccharide with the novel disaccharide repeat unit...polysaccharide synthesis requires a functional chromosomal *rfe* gene.

This paragraph also states that:

> Two LPS O-polysaccharide pathways have been described, and Rfe may play a role in either of these pathways. The O:54 O-polysaccharide shares a number of characteristics with the O-polysaccharide synthesized by the *Rfc*-independent O-polysaccharide biosynthetic pathway. (Emphasis added)

The above clearly describes sugar biosynthetic gene pathways.

Accordingly, Applicants respectfully submit that the claims have adequate written description and are enabled by the specification, and thus request withdrawal of the Examiner's rejection.

In paragraph 11, on page 4 of the Office Action, the Examiner rejects Claim 101 under 35 U.S.C. § 112, second paragraph.

Specifically, the Examiner contends that Claim 101 is indefinite with respect to which genes are considered to constitute "sugar-pathway genes".

As discussed above, "sugar-pathway genes" were well-known to one skilled in the art, and thus this expression is not indefinite.

Accordingly Applicants respectfully submit that the claims clearly and definitely recite the invention of interest. Thus, the Examiner is requested to withdraw the rejection.

In view of the arguments set forth above, reexamination, reconsideration and allowance are respectfully requested.

The Examiner is invited to contact the undersigned at his Washington telephone number on any questions which might arise.

Respectfully submitted,

Gordon Kit
Registration No. 30,764

SUGHRUE MION, PLLC
Telephone: (202) 293-7060
Facsimile: (202) 293-7860

WASHINGTON OFFICE
23373
CUSTOMER NUMBER

Date: January 16, 2004

THE JOURNAL OF BIOLOGICAL CHEMISTRY
1996 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 271, No. 45, Issue of November 8, pp. 28581–28592, 1996
Printed in U.S.A.

A Novel Pathway for O-Polysaccharide Biosynthesis in *Salmonella enterica* Serovar Borreze*

(Received for publication, July 10, 1996, and in revised form, August 20, 1996)

Wendy J. Keenleyside and Chris Whitfield‡

From the Department of Microbiology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

The plasmid-encoded gene cluster for O:54 O-polysaccharide synthesis in *Salmonella enterica* serovar Borreze ($rfb_{O:54}$) contains three genes that direct synthesis of a ManNAc homopolymer with alternating β1,3 and β1,4 linkages. In *Escherichia coli* K-12, $RfbA_{O:54}$ adds the first ManNAc residue to the Rfe (UDP-Glc*p*NAc::undecaprenylphosphate Glc*p*NAc-1-phosphate transferase)-modified lipopolysaccharide core. Hydrophobic cluster analysis of $RfbA_{O:54}$ indicates this protein belongs to the ExoU family of nonprocessive β-glycosyltransferases. Two putative catalytic residues and a potential substrate-binding motif were identified in $RfbA_{O:54}$. Topological analysis of $RfbB_{O:54}$ predicts four transmembrane domains and a large central cytoplasmic domain. The latter shares homology with a similar domain in the processive β-glycosyltransferases Cps3S of *Streptococcus pneumoniae* and HasA of *Streptococcus pyogenes*. Hydrophobic cluster analysis of $RfbB_{O:54}$ and Cps3S indicates both possess the structural features characteristic of the HasA family of processive β-glycosyltransferases. Four potential catalytic residues and a putative substrate-binding motif were identified in $RfbB_{O:54}$. In Δ*rfb E. coli* K-12, $RfbA_{O:54}$ and $RfbB_{O:54}$ direct synthesis of smooth O:54 lipopolysaccharide, indicating that this O-polysaccharide involves a novel pathway for O-antigen transport. Based on sequence and structural conservation, 15 new ExoU-related and 17 new HasA-related transferases were identified.

Lipopolysaccharide (LPS)[1] is a major component of the outer membrane of Gram-negative bacteria. The hydrophobic lipid A portion forms the outer leaflet of the outer membrane. In enteric bacteria, an O-polysaccharide is attached to lipid A via a short core oligosaccharide. The O-polysaccharide extends away from the cell surface to give a hydrophilic surface layer and represents a major surface antigen (O-antigen). It is characterized by a repeat unit structure, with different epitopes arising from variations in the nature, order, and linkage of the sugar monomers in the O-repeat unit and from the addition of side branches and "decorations" such as acetyl, ketal, and glycosyl residues.

Members of the genus *Salmonella* are serologically diverse. Approximately 50 O-antigens are recognized in *Salmonella*, and this diversity, combined with variable flagellar antigens, has resulted in the identification of greater than 2,000 serologically distinct *Salmonella* strains. Strains are assigned to serogroups based on shared expression of major O-antigen epitopes or factors (1). Synthesis of the O-antigen is directed by the products of the *rfb* gene cluster, and the structural diversity seen in the LPS O-antigens reflects variation in the *rfb* gene clusters; this variation is thought to have arisen from repeated lateral gene transfer and recombination events involving *rfb* genes. Selective pressure from the host immune response is proposed to be the driving force for the continued generation of antigenic diversity (2).

Serogroup O:54 is a heterogeneous group of 13 different *Salmonella* serovars. It is unique among *Salmonella* serogroups in that expression of the O:54 antigen requires the presence of a small plasmid. This is the only known plasmid-encoded O-polysaccharide in *Salmonella*. We have studied the 6.9-kb plasmid pWQ799 from serovar Borreze to determine the role of these plasmids in expression of the O:54 antigen (3). pWQ799 is a novel ColE1-related plasmid carrying the entire $rfb_{O:54}$ biosynthetic cluster (4). This is the only reported example of a ColE1-related plasmid carrying genes for the synthesis of cell-surface antigens. The plasmid is mobilized in the presence of an appropriate helper plasmid, providing the first defined mechanism for lateral gene transfer of O-antigen biosynthesis genes in *Salmonella enterica*. Mobilization of plasmids containing $rfb_{O:54}$ into strains possessing a chromosomally encoded *rfb* explains the simultaneous expression of two distinct O-antigens in members of this serogroup (3) as well as the serotyping results of plasmid-cured O:54 derivatives.

Enzymes encoded by the $rfb_{O:54}$ cluster direct the synthesis of an *N*-acetylmannosamine (ManNAc)-containing O-polysaccharide with the novel disaccharide repeat unit, →4)-β-D-Man*p*NAc-(1→3)-β-D-Man*p*NAc-(1→. O:54 polysaccharide synthesis requires a functional chromosomal *rfe* gene (3). Rfe transfers Glc*p*NAc-1-phosphate to undecaprenol-phosphate to form the first lipid intermediate (lipid I) in synthesis of the enterobacterial common antigen (5). Two LPS O-polysaccharide pathways have been described, and Rfe may play a role in either of these two pathways (reviewed in Ref. 6). The O:54 O-polysaccharide shares a number of characteristics with the O-polysaccharides synthesized by the Rfc-independent O-polysaccharide biosynthetic pathway including the following: 1) it is a homopolymer; and 2) Rfe is required for synthesis although GlcNAc does not form part of the repeating unit. The Rfc-independent pathway has so far only been identified for a few homopolymeric O-polysaccharides in *Klebsiella* O1 (7, 8), *Escherichia coli* O8 and O9 (9–11), and *Serratia marcescens* O16 (12). In this pathway, Rfe initiates synthesis by transferring Glc*p*NAc-1-P to the carrier lipid to form the acceptor upon which the complete O-polysaccharide chain is subsequently

* This research was supported by funding (to C. W.) from the Natural Sciences and Engineering Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) L39794.

‡ To whom correspondence should be addressed. Tel.: 519-824-4120 (ext. 3478); Fax: 519-837-1802; E-mail: cwhitfie@micro.uoguelph.ca.

[1] The abbreviations used are: LPS, lipopolysaccharide; ABC transporter, ATP-binding cassette transporter; ManNAc, *N*-acetylmannosamine; orf, open reading frame; HCA, hydrophobic cluster analysis; BLAST, basic local alignment search tool; ECA, enterobacterial common antigen; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; kb, kilobase pair(s).

polymerized. GlcpNAc therefore does not form part of the repeating unit. O-polysaccharides synthesized by this pathway are polymerized in the cytoplasm (13), and then the completed polymer is transported across the plasma membrane by an *rfb*-encoded dedicated ABC (ATP-binding cassette) transporter, prior to ligation to lipid A core. Analysis of $rfb_{O:54}$ and O:54 expression indicates that transport of the nascent O:54 polysaccharide occurs by a different mechanism. This communication describes the organization of $rfb_{O:54}$ and the characterization of the two glycosyltransferases that direct the synthesis of the O:54 O-polysaccharide. Our studies indicate that synthesis of O:54 LPS follows a novel mechanism distinct from either of the two characterized O-polysaccharide pathways.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Growth Conditions, and Plasmids—Bacterial strains were grown in Luria-Bertani broth supplemented, where appropriate, with antibiotics (ampicillin, 100 μg/ml; kanamycin, 50 μg/ml; chloramphenicol, 30 μg/ml; tetracycline, 15 μg/ml). For high level expression of $RfbA_{O:54}$ and $RfbB_{O:54}$, 100 mM isopropyl-β-D-galactopyranoside was added to LB broth cultures at a final concentration of 0.5 mM. Tn*phoA* mutants were examined on LB agar plates supplemented with 40 μg/ml 5-bromo-4-chloro-3-indolyl phosphate toluidine salt. The host for Tn*phoA* mutagenesis was *E. coli* CS118 (14), genotype Δ(*ara leu*) *7697 ΔlacX74 ΔphoA20 galE galK thi rpsE rpoB argE(Am) recA1*. LPS expression was examined in *E. coli* K-12 strains DH5α (15), Sϕ874 (16) or 21548 (17). The latter two strains have the following genotypes: Sϕ874, *lacZ trp ΔsbcB-rfb upp rel rpsL*; 21548, *thr-1 leuB6* Δ(*gpt-proA*)*66 hisG4 argE3 thi-1 rfbD1 lacY1 ara-14 galK2 xyl-5 mtl-1 mgl-51 rpsL31 kdgK51 supE44 rfe*::Tn*10*-48. The plasmids used in this study are shown in Fig. 1. The construction of plasmids pWQ800 and pWQ802 has been previously reported (3, 4). pWQ819 contains a 2.3-kb *Eco*RI-*Kpn*I fragment from pWQ799 cloned in pTrc99A (Pharmacia Biotech Inc.). pWQ820 was constructed by cloning the 2.4-kb *Mlu*1 fragment from pWQ799, after filling in the ends, into the *Sma*I site of pKK232-8 (Pharmacia), and then removing the insert using the flanking *Eco*RI and *Bam*HI sites from the vector multiple cloning region. This fragment was cloned into *Eco*RI-*Bam*HI-digested pGEM-7Zf(+) (Promega) to generate pWQ820. pWQ822 contains the 1.14-kb *Hin*cII-*Ssp*I fragment of pWQ799 inserted at the *Sma*I site of the expression vector pKK232-8. Plasmid pWQ823 was constructed by first cloning the 1.14-kb *Hin*cII-*Ssp*I fragment into the *Eco*RV site of pGEM-5Zf(+) (Promega) so that the direction of transcription was toward the SP6 promoter. The insert was then removed using the *Pst*I and *Nco*I sites of the vector and cloning this into *Nco*I-*Pst*I-digested pTrc99A.

DNA Manipulation and Analysis—Restriction enzyme digestions, ligations, and $CaCl_2$ transformations were performed as described by Sambrook *et al.* (15). DNA fragments to be subcloned were gel-purified using the GENECLEAN kit (Bio/Can Scientific, Mississauga, Ontario, Canada). Plasmids were column purified using QIAGEN spin columns (QIAGEN Inc., Chatsworth, Calif.) according to the manufacturer's instructions.

Nucleotide sequencing procedures are described elsewhere (4). Sequence data were edited and analyzed using AssemblyLIGN and MacVector software (International Biotechnologies Inc., New Haven, CT). Hydropathic analysis of the predicted protein sequence was done using the algorithm of Rao and Argos (18) with a minimum length for the transmembrane helix of 16. Homology searches of nucleotide and amino acid sequences in the National Center for Biotechnology Information data bases were done with the BLAST (basic local alignment search tool) server analysis program (19). Multiple alignments were performed using Clustal (20). Hydrophobic cluster analysis (HCA) was performed using the HCA-Plot program (Doriane Informatique, Le Chesnay, France). This program writes protein sequences on a duplicated α-helical net and circles clusters of hydrophobic amino acids (Ala, Val, Leu, Ile, Met, Phe, Gln). The plots are then visually compared for similarity in the hydrophobic cluster patterns, limiting analysis to the predicted globular portions of the proteins. β-Strands and α-helices are deduced based on the observed association of specific hydrophobic cluster shapes with secondary structures (21).

*Tn*phoA *Mutagenesis of Plasmid-encoded* $rfb_{O:54}$ *Determinants*—Plasmid-encoded $rfb_{O:54}$ genes were mutagenized using Tn*phoA* as described previously by Manoil and Beckwith (14). Fusions were made in plasmid pWQ800 (3). The precise site of insertion was mapped by sequencing out of the *phoA* orf using a primer (5′-CAGTAATATCGCCCTGAGCA-3′) that is complementary to the *phoA* sequence between nucleotides 79 and 98 (22).

RNA Isolation and Primer Extension—RNA was purified from a mid-log phase culture of *E. coli* DH5α (pWQ820) using TRIzol Reagent (Life Technologies, Inc.) according to the manufacturer's instructions. Primer extension reactions were done using the oligonucleotide primer TS154 (5′-TTTCATAATGTCGATCTGTTAATCC-3′) that corresponds to the complementary sequence to nucleotides 3676–3700 of plasmid pWQ799. The primer was end-labeled with [γ-^{32}P]ATP (DuPont NEN) and T4 polynucleotide kinase (Boehringer Mannheim, Laval, Quebec, Canada) and then purified using a QIAquick spin column (QIAGEN Inc., Chatsworth, CA). Primer extension experiments were performed using the First-Strand cDNA Synthesis Kit from Pharmacia (Pharmacia Biotech, Baie D'Urfé, Quebec) following the manufacturer's recommendations. DNA sequencing was done with the Sequenase version 2.0 sequencing kit (U. S. Biochemical Corp.).

LPS Extraction and Purification of Core Oligosaccharides—LPS samples were prepared either from SDS-proteinase K whole cell lysates, as described by Hitchcock and Brown (23), or by using a modification of the phenol/water extraction method; LPS was collected from both the aqueous and phenol phases (3). For compositional analysis of the LPS core oligosaccharides, lipid A was removed from phenol-purified LPS by hydrolysis in 1.5% acetic acid at 100 °C for 2 h. Precipitated lipid A was removed by centrifugation, and the supernatant was lyophilized. Samples were resuspended in water to a final concentration of 5 mg/ml, hydrolyzed, and analyzed by high performance anion-exchange chromatography as described previously (24).

Tricine-SDS-PAGE of LPS—LPS samples were analyzed by SDS-PAGE using commercially prepared 10–20% gradient Tricine gels from Novex (San Diego, CA). Electrophoresis conditions were those specified by the manufacturer. Gels were silver-stained using the method of Tsai and Frasch (25).

RESULTS

Organization of $rfb_{O:54}$—Previous analysis of the nucleotide sequence and genetic organization of pWQ799 indicated that approximately half of the 6915-base pair plasmid (nucleotides 72–3384) is involved in plasmid replication and mobilization (4). These regions are related to ColE1 and possess an average G + C content of 50–53% (Fig. 1*A*). This value is typical of *Salmonella* genomic DNA (26). In contrast, the remaining pWQ799 sequences have a uniformly lower average G + C content of 39%, with no detectable homology to any known ColE1-related sequences. The junctions between the high and low G + C regions in pWQ799 coincide with the 5′ and 3′ ends of the plasmid replicon regions. Abnormally low G + C values relative to those typical for the species is a common observation for genes involved in polysaccharide synthesis (27). These observations therefore suggested that the remaining DNA contained $rfb_{O:54}$ determinants.

Computer analysis for coding regions combined with sequence homology searches in the NCBI data bases (see below) resulted in the identification of three potential open reading frames (Fig. 1*B*; nucleotides 3547–4364, 4364–5740, and 5740–6873) designated $rfbA_{O:54}$, $rfbB_{O:54}$, and $rfbC_{O:54}$. Each gene is preceded by potential Shine-Dalgarno sequence: AGGA at nucleotides 3532–3535; AAGA at 4356–4359, and AGGA at 5727–5730. The initiation codons of $rfbB_{O:54}$ and $rfbC_{O:54}$ each overlap with the end of the preceding gene, suggesting that translation of the three genes is coupled. No consensus −10 and −35 *E. coli* promoter sequences were detected in the region immediately upstream of $rfbA_{O:54}$. To localize the $rfb_{O:54}$ promoter, the 1.14-kb *Hin*cII-*Ssp*I fragment of pWQ799, containing $rfbA_{O:54}$ and a 90-base pair 5′-flanking region, was cloned in front of the promoterless *cat* gene of pKK232–8. When transformed into *E. coli* K-12, the resulting plasmid (pWQ822; Fig. 1*C*) conferred chloramphenicol resistance, indicating the presence of an endogenous promoter. The transcription start site was mapped by primer extension of the oligonucleotide TS154 using total cellular RNA from *E. coli* DH5α (pWQ820; Fig. 1*C*). A single band was obtained in the resulting autoradiogram (data not shown), identifying the start site as nucleo-

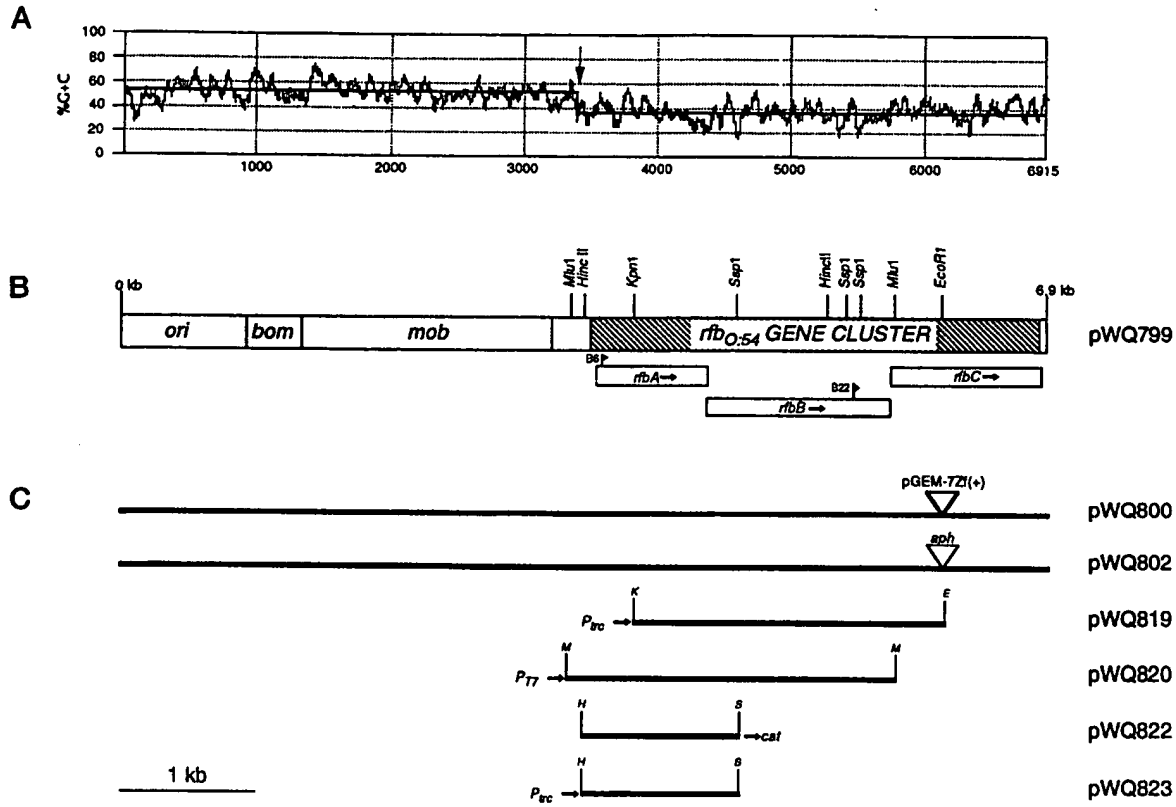

FIG. 1. **Structure and organization of pWQ799 and the *rfb*$_{O:54}$ cluster.** *A*, base composition plot of pWQ799. The average % G + C for the plasmid replicon sequences (nucleotides 1–3384) is 54%. The *arrow* indicates the beginning of *rfb*$_{O:54}$, where the % G + C drops to 39%. *B*, organization of pWQ799. The size, location, and direction of transcription of the *rfb*$_{O:54}$ orfs are indicated by the *open boxes* under the *stippled region* of the map. The functional regions of the pWQ799 replicon are presented for clarity; *ori* represents the origin of replication; *bom* is the origin of transfer; *mob* is the mobilization region. The *flags* represent the site of Tn*phoA* fusions; the *open flag* indicates PhoA$^-$, and the *filled flag* indicates PhoA$^+$. *C*, physical maps of plasmids used in this study. For simplicity, only the insert DNA is shown. The inserts in plasmids pWQ819 and pWQ823 are under the control of the *trc* promoter of vector pTrc99A. Expression of the insert in pWQ820 may be driven by the T7 promoter of vector pGEM-7Zf(+). The presence of the endogenous *rfb*$_{O:54}$ promoter in plasmid pWQ822 is indicated by expression of the promoterless *cat* gene in the vector pKK232–8. *aph* denotes the acetyltransferase gene that confers kanamycin resistance.

tide G at position 3474, 73 nucleotides upstream of the *rfbA*$_{O:54}$ start codon.

Previous studies of O:54 O-polysaccharide expression in strains containing pWQ800 (Fig. 1*C*) revealed a requirement for a functional *rffE* determinant in an *E. coli* host (3). This gene is part of the chromosomal enterobacterial common antigen (ECA) biosynthesis cluster and encodes the enzyme UDP-GlcNAc-2-epimerase (EC 5.1.3.14, UDP-*N*-acetyl-D-glucosamine 2-epimerase) (5). The epimerase converts UDP-GlcNAc to UDP-ManNAc, an intermediate in the formation of a precursor for ECA biosynthesis. UDP-ManNAc is also the presumed precursor for O:54 synthesis. Sequence analysis of pWQ799 indicated that the *Eco*RI site used in cloning of pWQ800 maps within *rfbC*$_{O:54}$. Subsequent mutant complementation analysis with an *rfbC*$_{O:54}$$^+$ subclone has shown that RfbC$_{O:54}$ is a functional homologue of RffE and is therefore not essential for O:54 synthesis in members of the Enterobacteriaceae, all of which synthesize ECA (28). The remaining two *rfb*$_{O:54}$ gene products, RfbA$_{O:54}$ and RfbB$_{O:54}$, are unique and essential for O:54 polysaccharide assembly and their structures and functions are described below.

Structural Analysis of RfbA$_{O:54}$—Sequence analysis of *rfbA*$_{O:54}$ predicts a 274 amino acid protein with a molecular mass of 31.8 kDa and a pI of 9.76. Hydropathic analyses (18) identified one transmembrane helix at the carboxyl terminus of the protein (residues 227–251). To determine the topology of RfbA$_{O:54}$, a promoterless leader sequence-deficient Tn*phoA* gene was inserted in-frame at a position corresponding to amino acid 21 of RfbA$_{O:54}$ (Fig. 1*B*). *E. coli* CC118 (pWQ800B6) was O:54-deficient and PhoA-negative on indicator media, indicating that this region of the protein is present in the cytoplasm. This cytoplasmic location is in agreement with the function of RfbA$_{O:54}$ (see below) and with protein topology predictions by the positive-inside rule. This rule allows the prediction of the topology of a bacterial inner membrane protein based on the observation that positively charged amino acids (Arg + Lys) are more abundant in cytoplasmic loops (*i.e.* ~15%) than in periplasmic loops (~5%) (29). The Arg + Lys content of this region was 14%.

The translated amino acid sequence of RfbA$_{O:54}$ shares significant homology with a number of putative bacterial glycosyltransferases (Fig. 2 and Table I). RfbA$_{O:54}$ shares 32% identity with the predicted product of the *rfb*$_{EcO7}$ gene orf275 (30), 26% identity with the hypothetical protein 6 of the *lsg* locus of *Haemophilus influenzae*, and 25% identity with the AmsE protein of *Erwinia amylovora* (31). Analysis of the protein alignments of these predicted products revealed that the sequence conservation was relatively uniform throughout the length of the proteins (Fig. 2). A number of protein sequences in the data bases were also identified with significant levels of homology over the N-terminal 192 amino acids of RfbA$_{O:54}$. Alignment of this region gave identity levels of 21 and 20%, respectively, with the ExoU and ExoO proteins from *Rhizobium meliloti* (32–34) and 21% with the LgtA glycosyltransferase from *Neis-*

```
RfbA    ---MGHQFTVCMVVTRNDAAEHFNVALLSLVCQSMKPSEVLIVDNGGLTDRHYEIIANIG
AmsE    ------MFSVLISLYNKEKPENLEQCLESLHQQTLNADEIVLVYDGPVSESLKAVATRWA
orf275  MSDDTPKFSVLMATYIKDSPLFLSEALQSIYKNTVAPDEVIIIRDGKVTSELNSVIDSWR
Lsg     -----MKFSVLMSLYIKENPQFLRECFESLVAQTRQADEIVLVFDGVVTPDLEFVVTEFE
                *.* . .   .   . *.  ..   *.... * ..    .

RfbA    DELSIRLLRIAHLSSVGQARNIALHAARYQVIAVADPDDINEPDRFIKLIPEIN--DSVR
AmsE    NLLPLVIVPLEKNLGLGKALNAGLERCTHNVVARMDTDDICLPERFEKQISYMESHPEVV
orf275  RYLNIKDFTLEKNMGLGAALNFGINQCMHDLVIRADSDDINRTNRFECILDFMTKNGDVH
Lsg     TKLPLKLVKLPQNRGLGKALNEGLLHCDYDWVFRMDTDDICVPDRFEKQVAFIEQHPESI
          * .   . .  .* * * *  .   .  * ***   **   .  .

RfbA    MVGAWVREFSREKNDRGIVRRVPCEFVDIIRYSRFRSPVNNPTICYFKNDALSVGGYNTS
AmsE    LSGAAVIEFDEHGKE--RLKRLPLSNNDIHEFARMKNPFNHMCVVFRKDKVISAGSYQHH
orf275  ILSSWVEEFEFNPGDKGIIKKVPSRNS-ILKYSKNRSPFNHPAVAFKKCEIMRVGGYGNE
Lsg     IFGGQIAEFGKNVNDIVAYRNVPTSAQEIIKFTQKRCPFNHMTVAYQKSAVINCGGYED-
          .  . **   .  . .*   *  ... . * *. ... *   .  * *

RfbA    LHFGEDYDLVTRFIRNGKIIRNIPIVMVNFRIGDAKKLYKKRGGLFLLKQEINLHRRFLE
AmsE    L-YMEDYNLWLRIMSLGHPVANLPDVLMKVRAGSDMVNK--RRGWNYIKSEVQLYRLKLA
orf275  Y-LYEDYALWLKSLANGCNGDNIQQVLVDMRPSKETAKR--RGGIKYAISEIKAQYHFYR
Lsg     --LQEDYYLWIKLVAQGLYMANLPDILVYARVGNGMVSR--RRGVNQAKAEWRLFKLKYR
          *** *  . .  *   *.  ...  *          * *      *

RfbA    NRYISLLDFFMIVTIKILVRLLPCWLFNVFYFKVLRKHE
AmsE    LKQTGFIRGTLYFLIRTMTRLMPVKVMQFLYEKDRKG--
orf275  ANYISYQDFIINIITRIFVRLLPTSFRGYTYKKVIRRFL
Lsg     LGIQGLLSGLFTFALRFGSRLLPTSLLKKLYQTFLRK--
                .  **.*      *
```

FIG. 2. **Multiple sequence alignment of $RfbA_{O:54}$ and homologous proteins.** Complete protein sequences are shown. References and accession numbers for the aligned proteins are listed in Table I. Identical amino acids are indicated by *asterisks*; similar amino acids are indicated by the *dots*.

TABLE I

Members of the ExoU family of nonprocessive β-glycosyltransferases

Protein	Putative or known activity	Organism	Accession no.	Ref.
$RfbA_{O:54}$	Adds first ManNAc residue during synthesis of the 0:54 O-polysaccharide	*S. enterica* sv. Borreze	L39794	This study
Orf275	Required for synthesis of the O-7 O-polysaccharide	*E. coli* O7	L04596	30
AmsE	Required for synthesis of amylovoran	*Erwinia amylovora*	X77921	31
AmsB	Required for synthesis of amylovoran	*E. amylovora*	X77921	31
Lsg6	Involved in LOS synthesis	*Haemophilus influenzae*	M94855	Unpublished
LgtA	Adds GlcNAc β1→3 to Gal; Involved in LOS synthesis	*Neisseria gonorrhoeae*	U14554	35
LgtD	Adds GalNAc β1→3 to Gal; involved in LOS synthesis	*N. gonorrhoeae*	U15992	35
Lsi-2	Adds GlcNAc β1→3 to Gal; involved in LOS synthesis; homologue of LgtA	*N. gonorrhoeae*	U15992	74
LgtD	Putative glycosyltransferase involved in LOS synthesis	*H. influenzae*	L46209, L42023	75
LgtD	Putative glycosyltransferase involved in LOS synthesis	*H. influenzae*	L45506, L42023	75
RfpA	Involved in synthesis of type 1 O-polysaccharide	*Shigella dysenteriae*	S73325	76
TrsC	Putative glycosyltransferase involved in outer core LPS synthesis	*Yersinia enterocolitica* O3	Z47767, X63827	77
GgaB	Involved in secondary teichoic acid synthesis	*Bacillus subtilis* 168	U13979	Unpublished
Orf4	Involved in staphyloxanthin synthesis	*Staphylococcus aureus* Newman	1340131	Unpublished
EpsI	Putative glycosyltransferase	*Streptococcus thermophilus* Sfi6	U40830	78
ExoU[a]	Adds Glc β1→6 to Glc; involved in succinoglycan synthesis	*R. meliloti*	L20758	33,34
ExoO[a]	Adds Glc β1→6 to Glc; involved in succinoglycan synthesis	*R. meliloti*	L20758	33,34
ExoA[a]	Adds Glc β1→3 to Gal; involved in succinoglycan synthesis	*R. meliloti*	L20758	33,34
ExoM[a]	Adds Glc β1→4 to acetylated Glc; involved in succinoglycan synthesis	*R. meliloti*	L20758	33,34
ExoW[a]	Adds Glc β1→3 to Gal; involved in succinoglycan synthesis	*R. meliloti*	L20758	33,34
Dpm1[a]	Dolichol mannose phosphate synthetase	*S. cerevisiae*	J04184	54
Orf2[a]	Unknown	*Anabaena* sp. strain PCC 7120	M31722	79

[a] Original member of the ExoU family, as defined by Saxena *et al.* (36).

seria gonorrhoeae (35) (Table I).

Previous studies of glycosyltransferases have demonstrated that there is often insufficient sequence similarity for functional predictions using traditional sequence alignments (36). However, transferases which catalyze the formation of glycosidic linkages with the same stereochemistry and with structurally related substrates are predicted to share a similar three-dimensional architecture in the catalytic and binding domains. This would be reflected in the presence of conserved structural regions or sequence motifs for shared mechanistic functions. Such domains can be identified using HCA (36, 37). This method plots the two-dimensional pattern of protein sequences and allows visual comparison and detection of conserved structural features. Using HCA, Saxena *et al.* (36) compared the two-dimensional structure of five different glycosyltransferases of known catalytic functions, including ExoO and ExoU from *Rhizobium meliloti*. They identified a structural region, domain A, that was present in all five proteins. The transferases all possess a common catalytic activity: formation of a single glycosidic linkage with a β-configuration from α-linked nucleotide diphospho sugar donors. Domain A is therefore speculated to be directly involved in this shared activity. Since the region of similarity detected between $RfbA_{O:54}$, ExoU, and ExoO includes the region containing domain A in the *Rhizobium* proteins, HCA was used to compare $RfbA_{O:54}$ with ExoU and the other $RfbA_{O:54}$ homologous proteins (Fig. 3). All of the proteins examined possessed the domain identified by Saxena *et al.* This domain is characterized by a series of vertical hydrophobic clusters typical of β-strands alternating with clusters characteristic of α-helices (36). Two conserved motifs were identified within the regions analyzed: (DXDD), at the C-terminal end of the β4 region of domain A, and EDY, which lies an average of 94 amino acids downstream of the (DXDD) motif, within another region of structural conservation. The

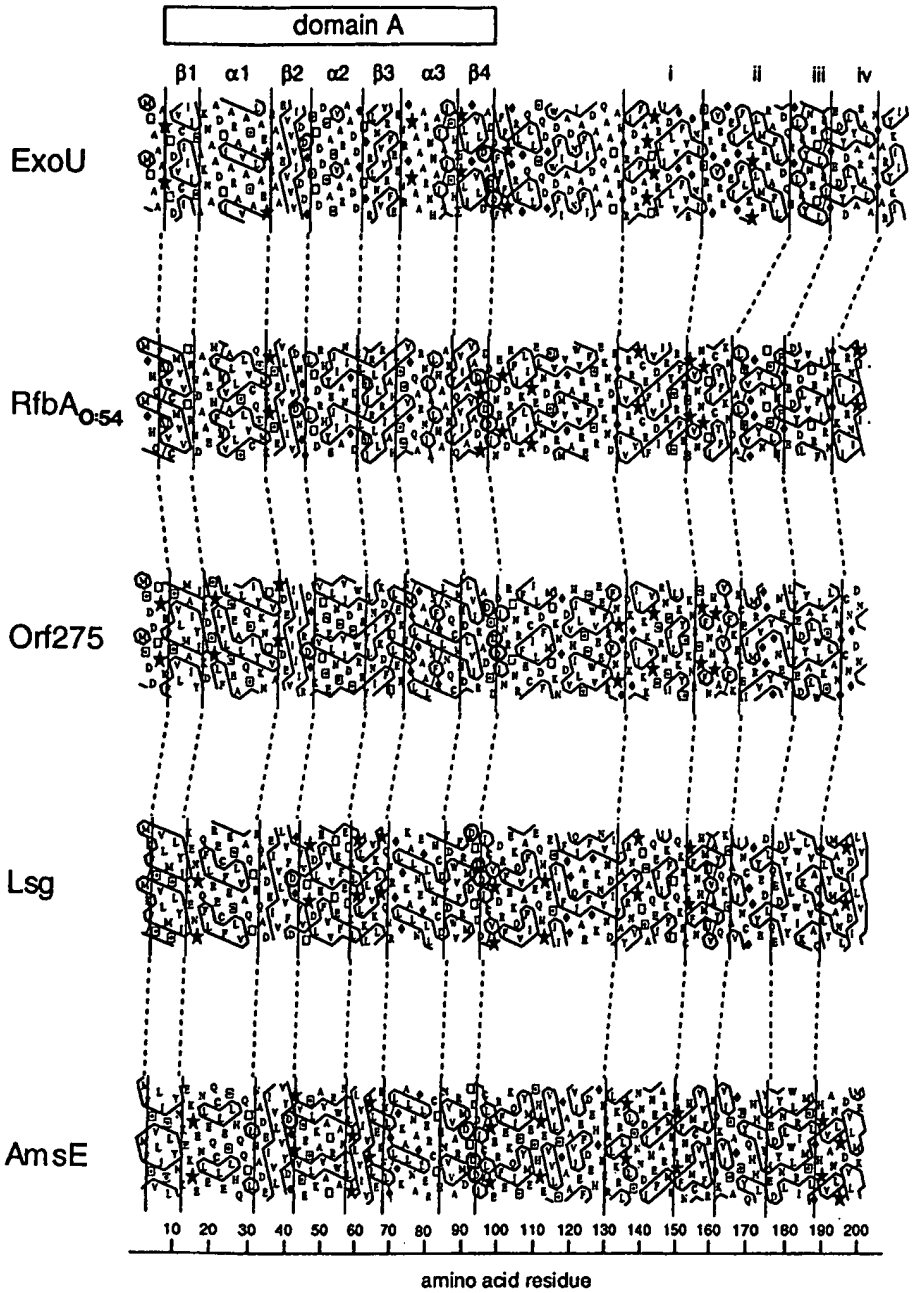

FIG. 3. **Alignment of HCA plots of ExoU, RfbA$_{O:54}$, and RfbA$_{O:54}$-related proteins.** Hydrophobic cluster analysis was performed using the HCA-Plot program (Doriane Informatique, Le Chesnay, France). This program writes protein sequences on a duplicated α-helical net and circles clusters of hydrophobic amino acids (Ala, Val, Leu, Ile, Met, Phe, Gln). The plots are then visually compared for similarity in the hydrophobic cluster patterns, limiting analysis to the predicted globular portions of the proteins. Plots were aligned using the results of amino acid sequence alignments as a starting point. Hydrophobic clusters with obvious similarities were used as anchors for the structural alignment, as were regions containing glycines (♦) and prolines (*), which are often present in loops (21). *Vertical lines* were drawn to indicate structurally conserved features. The prediction of β-strands and α-helices is based on the observed association of specific hydrophobic cluster shapes with secondary structures (21). Amino acid denoted with one-letter code except for proline (*), glycine (♦), serine (⊡), and threonine (□). Conserved residues are *circled*. Regions *i–iv* indicate structural regions that appear to have a conserved two-dimensional architecture surrounding the (EDY) motif; these regions were not described in the initial characterization of the ExoU family.

first motif includes a conserved Asp residue previously identified by Saxena's group. This residue falls in a loop following a β-sheet, and based on this conserved location and the acidic nature of the residue, Saxena *et al.* have speculated that this may be a catalytic residue. As noted by Saxena *et al.* (36), the proteins were also found to contain at least one Asp residue in the loop at the C-terminal end of the β2 strand (Fig. 3). The other protein sequences obtained from the data base search with RfbA$_{O:54}$ were also examined for the presence of these conserved features (Fig. 4). A total of 15 proteins, in addition to the 7 originally described by Saxena *et al.* (36), were identified as being members of the ExoU family (Table I).

RfbA$_{O:54}$ Is a Monofunctional N-Acetylmannosaminyltransferase—The observed sequence and structural homology of Rf-

```
                    Catalytic site 1                                                    Catalytic site 2
                             *                                                                 *    **
ExoU          36 A-EVVVID-DGSTDDSASVA          62 RLNVVRFEENRGPAAARNHAIA--ISHSPLIGVLDA-----DDFFFPGRLGQL
RfbA          34 PSEVLIV--DNGGLTDRHYE          63 RLLRIAHLSSV--GQARNIALH--AARYQVIAVADP-----DDINEPDRFIKL
ExoO          39 TVEVVVVD-DCSADATPALV           66 RLIALDRNRGP--GGARNAGIG--AARGRWIAVLDS-----DDTVRPDRLRRM
ExoM          37 RLRVIVADNDAEPSARALVE            66 ILYVHCPHSNI--SIARNCCLD--NSTGDFLAFLD------DDETVSGDWLTR
ExoW          32 DFHVLVID-DESPYPIADEL            60 RITVIRQPNGGP-GGARNTGLDNVPADSDFVAFLDS-----DDVWTPDHLLNA
ExoA          38 LNARVVIA-DGGSTDGTREI            66 RVLFLDNPKRIQ-SAAVNRAVAELGAGSDYLIRIDAHGTYPDDYCERLVEDAL
AmsE          31 ADEIVLVY-DGPVSESLKAV            60 VIVPLEKNLGL--GKALNAGLE--RCTHNVVARMDT-----DDICLPERFEKQ
orf275        37 PDEVIIIR-DGKVTSELNSV            66 KDFTLEKNMGL--GAALNFGLN--QCMHDLVIRADS-----DDINRTNRFECI
Lsg           32 ADEIVLVF-DGVVTPDLEFV            61 KLVKLPQNRGL--GKALNEGLL--HCDYDWVFRMDT-----DDICVPDRFEKQ

Dpm1          36 KTELIFVD-DNSQDGSVEEV            65 RIIVRTNERGL--SSAVLKGFY--EAKGQYLVCMDA------DLQHPPETVPK
```

```
                    Binding site 1
                             **
ExoU         190 LG----------------EDYDLYARALANGARYKIIHSCGYAAVVRGNSLSGSHRTIDLKRLY
RfbA         178 FG----------------EDYDLVTRFIRNGKIIRNIPIVMVNFRIG------DAKKLYKKRGG
ExoO         192 IG----------------EDYILLASALACGGALRGRAVRRLHLSYPRGFHLACVEARSHRRDD
ExoM         184 GG----------------EDTDFFTGMHCAGGTIAFSPEAWVHEPVPENRASLAWLAKRRFRSG
ExoW         176 IGRKLFEKVRFEATLKLAAEDV-LFFCDCVLASKRVVLCDAAGAVRGEGLNIFHSIDNDSPQFLK
ExoA         180 VG--------GYDESFSHNEDAELDYRLGKAGYRIWMTDKTSMVYYPRAKLVPLFWQYFGYGR-G
AmsE         166 AG---------SYQHHLYMEDYNLWLRIMSLGHPVANLPDVLMKVRAG--------SDMVNKRRG
orf275       173 VG---------GYGNEYLYEDYALWLKSLANGCNGDNIQQVLVDMRFS---------KETAKRRG
Lsg          169 CGG----------YEDLQEDYYLWIKLVAQGLYMANLPDILVYARVG--------NGMVSRRRG

Dpm1         168 YLENCNPRDINSQGFKIALELLAKLPLPRDPRVAIGEVPFTFGVRTEGESKLSGKVIIQYLQQLK
```

FIG. 4. **Multiple sequence alignment of conserved structural regions in ExoU, RfbA$_{O:54}$, and related proteins.** Conserved residues are indicated by *boldface*. *Shaded boxes* denote residues that are 100% conserved. Putative catalytic and binding sites are indicated *above* the alignment and were identified based on their highly conserved location and shared features with catalytic and binding sites of inverting glycosidases.

bA$_{O:54}$ with other glycosyltransferases suggested that *rfbA*$_{O:54}$ encodes a ManNAc transferase (EC 2.4.1.*X*). However, two separate ManNAc transferase activities are required for the generation of a ManNAc homopolymer with alternating β1→3 and β1→4 linkages. In other homopolysaccharide O-antigens where polymerization is initiated by Rfe, the transferase which adds the first residue to the Rfe intermediate is distinct from the enzymes which subsequently polymerize the O-polysaccharide (8, 9). A third transferase activity is therefore predicted for the transfer of the first ManNAc residue to undecaprenol-P-P-GlcNAc, the product of the Rfe reaction (5). To determine which *rfb*$_{O:54}$ gene product was responsible for the initial ManNAc transfer, plasmids carrying either *rfbA*$_{O:54}$ (pWQ823; Fig. 1*C*) or *rfbB*$_{O:54}$ (pWQ819; Fig. 1*C*) were transformed into different *E. coli* K-12 backgrounds, and LPS in the whole cell lysates of the transformants was analyzed by SDS-PAGE (Fig. 5). In *E. coli* DH5α (*rfe*$^+$*rffE*$^+$), neither gene was sufficient for O:54 synthesis and expression of a ladder of O-antigen-substituted LPS. Introduction of the two genes together on plasmid pWQ802, either in strain DH5α or the *rfb*-delete strain Sϕ874, was sufficient for expression of authentic O:54 O-polysaccharide. These results were confirmed by Western immunoblot using absorbed polyclonal O:54 antisera (data not shown). Analysis of the LPS core regions in SDS-PAGE profiles revealed that in the presence of *rfbA*$_{O:54}$, an additional LPS band was synthesized; this band migrated slightly slower than the lipid A core fraction of the host strain. In contrast, LPS from the *rfbB*$_{O:54}$-containing strain was indistinguishable from that of the parental strain. Further analysis determined that the RfbA$_{O:54}$-mediated band was not synthesized in an *rfe*$^-$ host strain, as would be expected for a biosynthetic pathway initiated by Rfe (Fig. 6*A*). The structure of the *E. coli* K-12 core oligosaccharide has been determined (38, 39) and does not include any ManNAc residues. The demonstration of ManNAc in the RfbA$_{O:54}$-modified core would therefore be indicative of a ManNAc transferase function catalyzed by RfbA$_{O:54}$. Purified

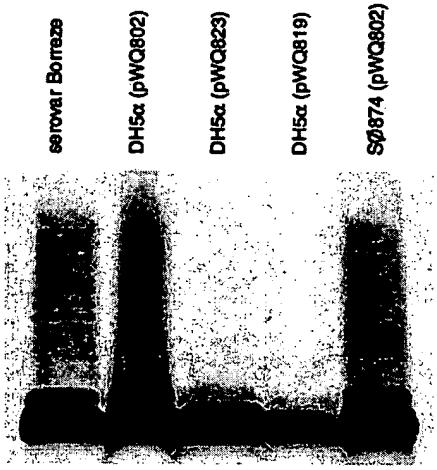

FIG. 5. **Analysis of the effect of RfbA$_{O:54}$ and RfbB$_{O:54}$ on LPS in whole cell lysates.** Silver-stained SDS-PAGE from LPS of strains containing either *rfbA*$_{O:54}$ (pWQ823), *rfbB*$_{O:54}$ (pWQ819), or *rfbA*$_{O:54}$ and *rfbB*$_{O:54}$ (pWQ802).

core oligosaccharide from *E. coli* DH5α (pWQ823) was therefore analyzed by high performance anion-exchange chromatography. A single additional peak corresponding to mannosamine, the acid-hydrolyzed product of ManNAc, was detected in the chromatogram of the RfbA$_{O:54}$-dependent core oligosaccharide (data not shown).

To determine the size of the RfbA$_{O:54}$-dependent band, the LPS from *E. coli* DH5α (pWQ823) was analyzed by SDS-PAGE alongside two LPS samples which each contain a modified core fraction of known size and composition (Fig. 6*B*). Plasmid pJK2363 contains the *Shigella dysenteriae* galactopyranosyltransferase gene, *rfpB* (40). Rfe$^+$ K-12 host strains expressing *rfpB* synthesize a core oligosaccharide modified by the addition

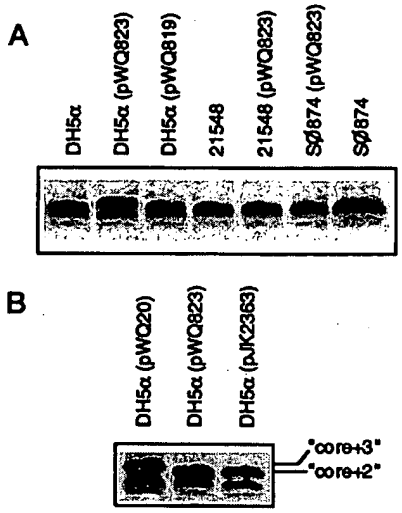

FIG. 6. **SDS-PAGE analysis of the $RfbA_{O:54}$-mediated core modification in whole cell lysates of different *E. coli* K-12 host strains.** *A*, analysis of the effect of host functions on $RfbA_{O:54}$-dependent core modification. Strains and plasmids are indicated *above* each lane. *B*, determination of the size of $RfbA_{O:54}$-modified core. In *E. coli* DH5α, plasmid pWQ20 directs the synthesis of a core fraction modified by the addition of three sugars (*core+3*); pJK2363 directs the formation of a core fraction modified by the addition of two sugars (*core+2*).

of the disaccharide [α-D-Gal*p*-(1→3)-β-D-Glc*p*NAc-(1→] (41). Plasmid pWQ20 contains the $rfbF_{KpO1}$ gene from *Klebsiella pneumoniae* O1, and *E. coli* K-12 strains containing $rfbF_{KpO1}$ produce a core oligosaccharide modified by the addition of the trisaccharide [β-D-Gal*f* -(1→3)-α-D-Gal*p*-(1→3)-β-D-Glc*p*NAc-(1→] (8). The comigration of the $RfbA_{O:54}$-modified band with the "core + 2" band of *E. coli* K-12 (pJK2363), combined with the Rfe dependence of this band and chemical compositional analysis all indicate that the $RfbA_{O:54}$-dependent band represents lipid A core modified by the addition of a single GlcNAc and a single ManNAc residue.

The $RfbA_{O:54}$-modified lipid A core fraction was not present in *E. coli* Sφ874 (pWQ823), indicating that synthesis of this LPS fraction requires one or more function(s) encoded by the K-12 *rfb* gene cluster (Fig. 6*A*). Expression of O:54 LPS in Sφ874 (pWQ802) indicates that this requirement is overcome by the simultaneous activities of $rfbA_{O:54}$ and $rfbB_{O:54}$. Ligation of newly polymerized O-chains to lipid A core occurs at the periplasmic face of the plasma membrane and must therefore follow trans-plasma membrane transport of the undecaprenol-bound intermediate. A similar $rfb_{K\text{-}12}$-dependent transport activity has been demonstrated for the $RfbF_{KpO1}$ and RfpB-mediated core modifications (8, 42). The *E. coli* K-12 O-antigen transporter, RfbX, is believed to be responsible for the transport of the lipid-bound $RfbF_{KpO1}$ and RfpB products (8, 42, 43). Given the nature of this transport event and the involvement of both host and cloned plasmid functions, it was possible that the appearance of a single modified core band in *E. coli* DH5α (pWQ823) resulted from a substrate size limitation imposed by RfbX, rather than the monofunctional transferase activity of $RfbA_{O:54}$. In this case, lipid-linked intermediates with higher degrees of polymerization would be synthesized but would remain in the cytoplasm, attached to carrier lipid. To address this possibility, the O-polysaccharide ligase-deficient strain *E. coli* CS2334 was transformed with pWQ823. Under such conditions any O-polysaccharide that is formed, but not transported, should accumulate in the cytoplasm. *E. coli* CS2334 (pWQ823) was phenol-extracted, and because the polymeric O:54 LPS partitions primarily into the organic phase (3), both phases were examined. The linkage between undecaprenol and carbohydrate polymer is phenol-labile (43) and, as a consequence, extracted O-haptenic material remains in the supernatant following a 100,000 × *g* centrifugation step. Supernatants from both the aqueous and phenol phases were size-fractionated on a Sephadex G-50 column, and fractions containing amino sugars were analyzed by ^1H NMR. Although extracted high molecular weight ECA was detected in these experiments, no ManNAc-containing polymer was present (data not shown). These data confirm that $RfbA_{O:54}$ transfers a single ManNAc residue.

Sequence Analysis of $RfbB_{O:54}$—$RfbB_{O:54}$ is predicted to be a 53.3-kDa protein composed of 459 amino acids, with a calculated pI of 8.32. Hydropathic analysis predicted four transmembrane helices between residues 11–40, 325–340, 385–406, and 416–438. This protein is therefore expected to be an integral membrane protein. Based on the relative distribution of positive amino acid residues, the hydrophilic region defined by residues 340 and 384 is predicted to lie in the periplasm (6% Arg + Lys). This location was confirmed by construction of an in-frame PhoA-positive fusion (pWQ800B22) at amino acid 368 (Fig. 1*B*).

The demonstrated activity of $RfbA_{O:54}$ combined with the observation that $RfbA_{O:54}$ and $RfbB_{O:54}$ together are sufficient for synthesis of polymeric O:54 LPS suggests that $RfbB_{O:54}$ is a processive β-glycosyltransferase (EC 2.4.1.*X*). Searches of the data bases with the $RfbB_{O:54}$ sequence identified only one protein, IcaA (44), with significant sequence homology over its entire length (22%). IcaA is a putative glycosyltransferase from the *ica* (intercellular adhesion) locus of *Staphylococcus epidermidis*. This protein is believed to play a role in the synthesis of the linear β-1,6-linked glucosaminoglycan involved in intercellular adhesion. The predicted size and topology of this protein resembles that of $RfbB_{O:54}$. A number of other proteins were identified that possess regions of conservation with the large central hydrophilic domain of $RfbB_{O:54}$ (Table II). Most of these proteins are glycosyltransferases involved in the synthesis of bacterial cell surface polysaccharides. Two of these proteins, Cps3S and HasA, are also predicted to be integral membrane proteins with similar hydropathy plots and four predicted transmembrane domains. Cps3S is a glycosyltransferase that directs the synthesis of the type 3 capsule of *Streptococcus pneumoniae* (45, 46). The enzyme is bifunctional and processive, catalyzing the formation of [→3)-β-D-GlcA-(1→4)-β-D-Glc-(1→]$_n$ (45). HasA is the hyaluronic acid synthase from Group A *Streptococcus pyogenes* (47, 48). This protein is also a bifunctional processive β-glycosyltransferase, catalyzing the formation of a polysaccharide with the structure [→4)-β-D-GlcA-(1→3)-β-D-Glc-(1→]$_n$ (47).

Using HCA, Saxena *et al.* (36) compared the plots of HasA with a number of known processive β-glycosyltransferases. They reported a correlation between the presence of two conserved structural regions (domains A and B) and the shared catalytic activity. Domain A is common to both the ExoU and HasA families of β-glycosyltransferases; whereas domain B, located a short distance downstream of domain A, is unique to the HasA family. A single conserved Asp residue in region II and a conserved sequence motif (Q*X*XRW) in region IV were both reported to characterize domain B. Comparison of the HCA plots of the large hydrophilic domains of $RfbB_{O:54}$, Cps3S, and IcaA, with that of HasA, confirmed the presence of both domains in these proteins (Fig. 7). All three proteins possessed the conserved Asp and (Q*X*XRW) motif of domain B and with the exception of Cps3S, the two conserved Asp residues characteristic of domain A were also present in $RfbB_{O:54}$ and IcaA (Fig. 7). Although the Asp-β4 is conserved in Cps3S, there is no

TABLE II
Members of the HasA family of processive β-glycosyltransferases.

Protein	Putative or known function	Organism	Accession no.	Reference
$RfbB_{O:54}$	Polymerizes O:54 O-polysaccharide; transfers ManNAc to ManNAc in alternating β1→3 and β1→4 linkages	*S. enterica* sv. Borreze	L39794	This study
Cps3S	Cellubiuronic acid synthase; synthesizes $[\rightarrow 3)$-β-D-GlcA-$(1\rightarrow 4)$-β-D-Glc-$(1\rightarrow 1]_n$	*S. pneumoniae*	Z47210	46
			U15171	45
IcaA	Involved in intercellular adhesion	*S. epidermidis*	U43366	44
HmsR	Membrane protein involved in autoagglutination	*Y. pestis*	U22837	Unpublished
NodC	Polymerizes Nod factor oligosaccharide; transfers GlcNAc β1→4 to GlcNAc	*Rhizobium loti*	X52958	80
NodC	Nod factor oligosaccharide synthesis	*R. galegae*	X87578	Unpublished
NodC	Nod factor oligosaccharide synthesis	*R. meliloti*	X01649	81
			M11268	82
NodC	Nod factor oligosaccharide synthesis	*R. leguminosarum*	Y00548	83
NodC	Nod factor oligosaccharide synthesis	*R. leguminosarum* bv. phaseoli	M58626	84
NodC[a]	Nod factor oligosaccharide synthesis	*Azorhizobium caulinodans*	L18897	85,86
FbfA	Involved in fruiting body morphogenesis; related to NodC proteins	*Stigmatella aurantiaca*	Z11601	Unpublished
Chs2	Chitin synthase; transfers GlcNAC β(1→4) to GlcNAc	*Candida albicans*	M82937	65
Cal1	Chitin synthase	*S. cerevisiae*	X57300	68
Chs3	Chitin synthase	*C. albicans*	D13454	67
Csd2	Chitin synthase	*S. cerevisiae*	M73697	64
ChsE	Chitin synthase	*Emericella nidulans*	U52362	Unpublished
Chs-4	Chitin synthase	*N. crassa*	U25097	Unpublished
CelA	Cellulose synthase; transfers Glc β1→4 to Glc	*Agrobacterium tumefaciens*	L38609	60
HasA[a]	Hyaluronic acid synthase; synthesizes $[\rightarrow 4)$-β-GlcA-$(1\rightarrow 3)$-β-GlcNAc-$(1\rightarrow]_n$	*Streptococcus pyogenes*	L21187	47,48
Alg8[a]	Involved in alginate synthesis; exact function unknown	*Pseudomonas aeruginosa*	L22611	62
Chs1[a]	Chitin synthase	*S. cerevisiae*	M14045	63
AcsAB[a]	Cellulose synthase	*Acetobacter xylinum*	X54676	53,59
Dg42[a]	Function unknown	*Xenopus laevis*	M22249	87

[a] Original member of the HasA family, as defined by Saxena *et al.* (36).

Asp immediately next to the β2 sheet of the protein; however, an Asp residue is located in the middle of α2 (Fig. 8). A search of the data bases identified a total of 17 other proteins, in addition to the 6 originally identified by Saxena *et al.*, that possess the conserved features of the HasA family (Table II).

DISCUSSION

Plasmid pWQ799 carries the entire O:54 biosynthetic cluster and consists of three genes, $rfbA_{O:54}$, $rfbB_{O:54}$, and $rfbC_{O:54}$. Together these genes direct the synthesis of a unique *Salmonella* O-polysaccharide. The O:54 polysaccharide is the only known homopolymeric O-antigen in *Salmonella* (49), and while the presence of *N*-acetylated amino sugars is not an unusual feature for *Salmonella* O-polysaccharides, ManNAc is found only in O:54. It is therefore not surprising that the predicted products of $rfb_{O:54}$ showed no homology to other *Salmonella* protein sequences in the data bases.

$RfbA_{O:54}$ is an *N*-acetylmannosaminyltransferase. This enzyme transfers the first ManNAc residue to undecaprenolpyrophosphoryl-GlcNAc, the product of the Rfe reaction. The involvement of Rfe (UDP-GlcpNAc::undecaprenolphosphate GlcpNAc-1-phosphate transferase) in the initiation of O:54 biosynthesis resembles the situation observed with other homopolysaccharide O-antigens (7, 10, 12, 50), where biosynthesis is initiated by Rfe but GlcNAc does not form part of the O-unit. It is unclear if this type of initiation reaction is necessarily limited to O-polysaccharides. $RfbA_{O:54}$ shares significant sequence homology with a number of glycosyltransferases in the data bases. Two of the proteins, ExoO and ExoU from *R. meliloti*, are β-glycosyltransferases that catalyze β-1,6 linkages from α-linked nucleotide diphospho sugar donors. Alignment of the HCA plots derived for $RfbA_{O:54}$ and 4 $RfbA_{O:54}$-related proteins with the previously characterized plot of ExoU (36) revealed that all of the proteins possess the same N-terminal structural region (domain A) identified by Saxena and co-workers (36) in ExoU and a number of functionally related transferases. This domain is believed to be responsible for the nonprocessive β-glycosyltransferase activity of proteins within this family. Given their common activities and the conservation of domain A among the members of the ExoU family of glycosyltransferases, it is likely that this region represents at least part of the catalytic domain of these enzymes. We identified two highly conserved motifs among the aligned HCA plots. The first (DXDD),falls in the β4 region of domain A and includes a conserved Asp residue previously identified by Saxena and co-workers (36). Saxena *et al.* (36) also identified a second conserved Asp residue in domain A, in the loop at the C-terminal end of the β2 strand of domain A. HCA alignments revealed that all of the proteins possess either one ($RfbA_{O:54}$, orf275, Lsg6, and AmsE) or two Asp residues (ExoU) in this region, although the position is not strictly conserved and an Asp may occupy the first and/or the second site within this loop.

Very little is known about anabolic glycosyl transfer reactions, but it has been proposed that, mechanistically, this type of reaction may be viewed as the reverse reaction of the glycosyl transfer reaction performed by O-glycosidases, the difference being that the result is the extension rather than hydrolysis of an oligosaccharide or polysaccharide chain (51). By analogy with the extensively characterized polysaccharide hydrolase systems (reviewed in Refs. 51, 52), this hypothesis predicts that formation of a β-glycosyl linkage from an α-linked sugar nucleotide donor would involve the same type of catalytic event as that of the inverting glycoside hydrolases. Hydrolysis of glycosidic bonds by inverting glycosidases results in a net inversion of configuration at the anomeric center of the reducing sugar product. The catalytic mechanism involves two acidic active site amino acids that act as acid-base catalysts. Among the cellulases and xylanases, in every instance where the catalytic residues have been identified, the amino acid has been either an aspartate or a glutamate (52). One of these residues is believed to act as the acid catalyst to protonate the substrate, whereas the other is thought to act as the base catalyst by deprotonating water. The two catalytic residues are located in flexible loop regions in the active site cleft, between substrate-binding subsites. Characterization of $RfbA_{O:54}$-related proteins in data bases has identified 15 new members of the ExoU family, in addition to the 7 originally described by Saxena *et al.*

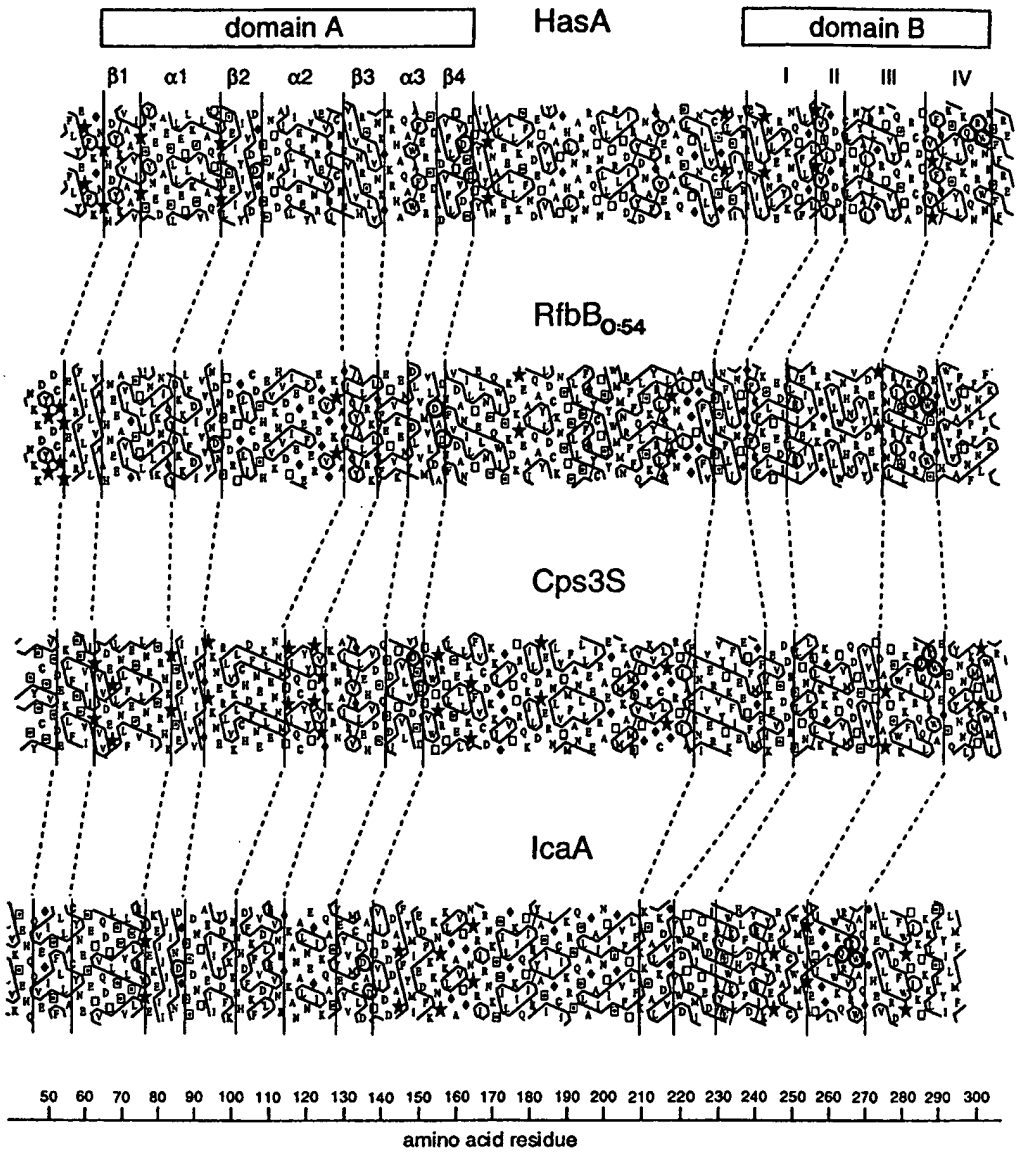

FIG. 7. **Alignment of the HCA plot of HasA with plots of RfbB$_{O:54}$ and related proteins.** Plots were generated using the method outlined for Fig. 3. Conserved residues are *circled*.

(53). These new members all possess two conserved Asp residues within the putative catalytic domain (Fig. 4). The second, in the terminal loop of β4, falls in a more strictly conserved region of the proteins and is surrounded by additional acidic residues. In RfbA$_{O:54}$, sequence and HCA plot alignments predict that the catalytic residues are Asp-41 and either Asp-93, Asp-95, or Asp-96.

HCA also identified a conserved motif (EDY) downstream of domain A, in a region containing amino acid clusters with a shape typical of α-helices (21). Carbohydrate ligands are believed to interact with binding subsites through a combination of ionic interactions, hydrogen bonding, and stacking interactions between aromatic amino acids and the hydrophobic patches of carbohydrate monomers (reviewed in Ref. 52). The conservation of the (EDY) motif and its position within a region of structural conservation, a short distance from a putative catalytic residue, suggests this motif might represent part of a binding subsite. Multiple sequence alignment revealed that a similar motif occurs in other members of this family (Fig. 4) although the order of the first two acidic amino acids may be reversed and in some cases the third amino acid may be a Lys. Although Tyr is an aromatic amino acid, both it and Lys are structurally similar, in that both have long chains and may participate in hydrogen bonding through either a terminal hydroxyl (Tyr) or amino (Lys) group. The acidic nature of Asp and Glu would also allow hydrogen bonding of these residues with a carbohydrate ligand. Given the predominance of aromatic amino acids surrounding the (EDY) motif, it is possible that the surrounding residues participate in stacking interactions with the hydrophobic patch of the glycosyl monomer. The only protein that did not possess the (EDY) motif was Dpm1. This enzyme catalyzes the transfer of mannose from GDP-man to dolichol phosphate (54). The fact that the ligand recognized by Dpm1 is not a carbohydrate is further support for the hypothesis that this motif is involved in substrate binding.

RfbB$_{O:54}$ is predicted to be an integral membrane protein with four transmembrane helices, a large central hydrophilic domain and a periplasmic loop. Searches of the data bases

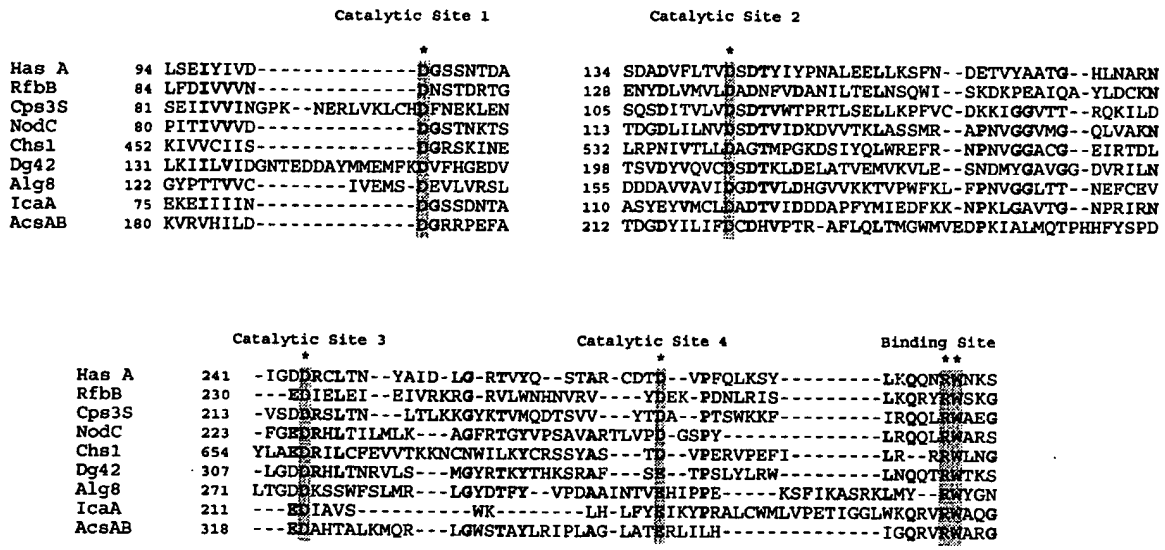

FIG. 8. **Multiple sequence alignment of conserved structural regions in HasA, $RfbB_{O:54}$, and related proteins.** Conserved residues are indicated by *boldface*. Residues that are 100% conserved are indicated by *shaded boxes*. Putative catalytic and binding subsites were identified by extrapolation from the inverting glycosidases.

identified a number of proteins possessing regions of conservation with the central hydrophilic domain. Among these, the proteins with known activities were all glycosyltransferases involved in synthesis of cell surface or secreted polysaccharides. The demonstrated activity of $RfbA_{O:54}$ combined with the fact that *rfbAB*$_{O:54}$ are sufficient for polymeric O:54 LPS in an *E. coli* Δ*rfb* host suggest that $RfbB_{O:54}$ is a processive ManNAc transferase that acts after $RfbA_{O:54}$ to polymerize the O:54 polysaccharide. Analysis of the sequence alignments obtained with $RfbB_{O:54}$ identified regions of homology with two known processive β-glycosyltransferases, Cps3S from *S. pneumoniae* (45, 46) and HasA from Group A *S. pyogenes* (47, 48). Both proteins have four predicted transmembrane domains with a large central hydrophilic region, and HCA of these hydrophilic domains revealed that $RfbB_{O:54}$ and Cps3S both possess the same multidomain structure and conserved residues previously identified in the HasA family of processive β-glycosyltransferases recently described by Saxena *et al.* (53). There are two conserved structural regions in proteins within this family. The first region (domain A) is common to both the ExoU and HasA families, and the second (domain B) is specific to the HasA family. Domain B is subdivided into four regions and has a conserved Asp residue at the C-terminal end of region II and a conserved sequence motif (QXXRW) at the C-terminal end of region IV. A search of the data bases for proteins possessing the conserved features of this family of proteins identified 17 new members (Table II), including $RfbB_{O:54}$, Cps3S, and IcaA. Of the proteins that were identified, most are known processive enzymes, including Cps3S; five are NodC proteins involved in the polymerization of the β-1,4-linked Nod factors of Rhizobia; eight are chitin synthases, involved in the polymerization of β-1,4-linked GlcNAc polysaccharide chains of yeast chitin; and one is a cellulose synthase, involved in the polymerization of the β-1,4-linked Glc chains of cellulose. The identification of $RfbB_{O:54}$ as a member of this family provides further evidence for the processive nature of this enzyme; however, it must be acknowledged that the data do not preclude the possible involvement of $RfbA_{O:54}$ in the elongation process.

It is interesting to note that synthesis of the *E. coli* K5 capsule also involves a processive glycosyltransferase, KfiC (55). In this polysaccharide, the repeating unit consists of alternating β- and α-linkages, [→4)-β-D-Glc*p*A-(1→4)-α-D-Glc*p*NAc-(1-] (56). KfiC must therefore catalyze both α- and β-linkages from α-linked donor molecules. Sequence alignment between $RfbB_{O:54}$ and KfiC revealed that the K5 transferase possesses domain A but not domain B (data not shown), providing further support for the role of each domain in the formation of β-glycosidic bonds.

Saxena *et al.* (36) have proposed a model for the processive mechanism of polymerization in the HasA family of proteins. The model accounts for the characteristic multidomain architecture of the HasA family of proteins and, as with the monofunctional β-glycosyltransferases, is based on the hypothesis that inverting anabolic glycosyltransferases use the same catalytic mechanism as inverting glycoside hydrolases (51). According to this model, domains A and B represent different catalytic domains that together allow these proteins to catalyze two β-glycosidic bonds, either simultaneously or sequentially. The subsequent loss of the two UDP groups from the catalytic sites is proposed to provide the driving force for the chain to move through the catalytic cleft until the terminal sugar interacts with the last binding subsite, allowing two more UDP sugars to enter. The simultaneous formation of two glycosidic linkages provides a simple mechanism for the generation of the 2-fold screw axis that arises from a disaccharide repeat with two β-glycosidic linkages, without invoking a concomitant rotation of either the enzyme or the substrate. It also provides an effective mechanism for maintaining the fidelity of a heteropolysaccharide disaccharide repeat. However, by analogy with the inverting glycosidases, such an activity would involve a total of four conserved acidic amino acid residues, two in each domain. Sequence and HCA plot alignments identified only one conserved acidic residue in domain B. Based on sequence conservation in the chitin synthases, Nagahashi *et al.* (57) have also speculated that this residue is involved in catalysis. To further substantiate this hypothesis, these workers used site-specific mutagenesis to replace this conserved Asp (Asp-562) with the longer Glu and observed a 100% reduction in enzyme activity. They also replaced Asp-562 with an Asn, to determine whether the hydroxyl group of Asp was necessary for activity. Not surprisingly, a complete loss of activity was observed in the mutant. Because the inverting mechanism predicts two catalytic residues in domain B, the sequences of all of the known transferases in the HasA family were examined for additional

conserved amino acids. The search was limited to the region extending from the end of β4 to approximately 25 residues past the (QXXRW) motif. The C-terminal end of the search region corresponds to the start of a predicted transmembrane domain in $RfbB_{O:54}$, Cps3S, HasA, IcaA, and NodC. A highly conserved proline was identified a short distance in front of the (QXXRW) motif, in the junction between regions III and IV (Fig. 8). This Pro is predicted to lie within a loop at the C-terminal end of a β-sheet. In all of the proteins, an Asp or Glu was found 2–4 residues before the conserved Pro. It is possible that the conserved carboxylate adjacent to this Pro represents the second catalytic residue of domain B. In $RfbB_{O:54}$, the catalytic residues are therefore speculated to be Asp-92 or Asp-94, Asp-151, Asp-244, and either Glu-268 or Asp-269.

The position of the (QXXRW) motif relative to the fourth potential catalytic residue in these proteins suggests that the motif may represent part of a binding subsite. This possibility is supported by the predicted interactions of the residues in the motif with a carbohydrate ligand: hydrophobic interactions between the aromatic Trp residue and the hydrophobic patch of a glycosyl monomer, and hydrogen-bonding interactions between the guanidinium side chain of Arg and the glycosidic hydroxyls. Noting that this motif was highly conserved in the chitin synthases, Nagahashi *et al.* (57) also speculated on the possible function of this motif in enzyme activity. To confirm a structure-function relationship, they used site-specific mutagenesis to individually replace each residue in the motif, and then measured enzyme activity and K_m values in the resulting mutant proteins. In each case, a conservative change resulted in a reduction in activity and either an increase or decrease in binding affinity. While these results were interpreted as evidence for a role in catalysis, they could equally be interpreted as evidence for a role in binding of the substrate or in hydrogen-bonding interactions with a catalytic residue, thereby maintaining the correct orientation for catalysis.

There are presently two known pathways for O-antigen biosynthesis. These pathways are fundamentally different. Key criteria distinguishing the two are the cellular location of the polymerization reaction and the mode of export across the plasma membrane (6). In Rfc-dependent synthesis, polymerization involves block-wise addition of single O-repeat units and occurs at the periplasmic face of the plasma membrane. Individual O-units are assembled on undecaprenol-P in the cytoplasm and then transported across the plasma membrane, presumably by the O-unit transporter, Rfb*X* (58). Ligation to lipid A core is catalyzed by RfaL. Rfc-independent O-polysaccharide biosynthesis is currently limited to homopolymeric O-polysaccharides. In this pathway, synthesis is initiated by the Rfe-dependent transfer of GlcNAc-1-P to undecaprenol-P. The complete O-polysaccharide chain is then polymerized in the cytoplasm prior to being delivered to the site of ligation by an *rfb*-encoded dedicated ABC transporter. At first glance, the pathway for O:54 polysaccharide synthesis appears similar to the Rfc-independent pathway of O-antigen synthesis. In both cases, synthesis is initiated by Rfe and the O-repeat is a homopolymer. In addition, by analogy with other more well-characterized HasA-related proteins, $RfbB_{O:54}$ is expected to polymerize the O:54 polysaccharide in the cytoplasm, where pools of activated precursor are available. It is at this point in the pathway that O:54 biosynthesis diverges from the Rfc-independent pathway; there is no dedicated ABC transporter for export of the polymerized O:54 polysaccharide. This is the only known O-polysaccharide system that does not encode either an ABC transporter or an Rfb*X* O-unit transporter. Despite the absence of either of these components, smooth O:54 LPS is expressed in a Δ*rfb E. coli* K-12 host strain containing only $rfbA_{O:54}$ and $rfbB_{O:54}$. We have previously shown that, in the absence of the cognate ABC transporter, an *E. coli* K-12 strain containing the remaining *rfb* genes from *K. pneumoniae* O1 (Rfc-independent) accumulates O-antigen in the cytoplasm (13). There is, therefore, no alternate, generic O-polysaccharide export system in *E. coli*. Consequently O:54 synthesis represents a new pathway for O-antigen assembly, involving a different mechanism for delivering nascent O-polysaccharide to the LPS O-antigen ligase.

Intriguingly, no export system has yet been identified for polymers produced by a number of other processive β-glycosyltransferases in the HasA family. These polysaccharides include bacterial cellulose (*Acetobacter xylinum* and *Agrobacterium tumifaciens* (53, 59, 60)), hyaluronic acid (*S. pyogenes* (47, 48, 61), the type 3 capsule of *S. pneumoniae* (45, 46), alginate (*P. aeruginosa* (62)), and chitin (*S. cerevisiae*, *Candida albicans*, *Emericella nidulans* and *Neurospora crassa* (63–68)). In contrast, transmembrane transport of the rhizobial Nod factors occurs through the action of a dedicated ABC transporter (69, 70). However, synthesis of Nod signal factor differs somewhat from that of cell-surface polysaccharides, as a carrier lipid has not yet been identified; the polymer that is synthesized is much shorter and is generally thought to be secreted (71); and the oligosaccharide product is substituted with acyl, acetyl, and sulfate groups in a strain-specific manner (72). The NodC proteins are all highly homologous, and their predicted topology differs from that of $RfbB_{O:54}$, HasA, and Cps3S. NodC proteins may therefore represent a subfamily of the HasA family. The predicted topology of $RfbB_{O:54}$, with a periplasmic loop following the cytoplasmic glycosyltransferase domain, suggests the possibility that the protein possesses two separate activities, catalyzing the polymerization of the O:54 polysaccharide and coupling this with transport in a vectorial reaction, similar to that suggested for chitin synthases (73). The C-terminal transmembrane domains would therefore be predicted to form a pore or channel through which the growing chain is extruded. Similarities in size, two-dimensional architecture, and hydropathy plots point to the possibility of a similar transferase/transport function for Cps3S, IcaA, and HasA.

Synthesis of the O:54 polysaccharide clearly represents a third and new pathway for O-antigen assembly. With the identification of potential catalytic residues in $RfbB_{O:54}$ and the speculated role for the periplasmic loop in transmembrane transport, the O:54 O-polysaccharide provides a relatively simple system for testing this putative export function and for examining the mechanism of catalysis of β-glycosyltransferases. These analyses will also serve to characterize a novel O-polysaccharide biosynthetic pathway.

Acknowledgments—We thank Dr. A. J. Clarke for his help in the analysis and presentation of the multiple sequence alignments and for critically reviewing the manuscript. We also thank Malcolm Perry at the National Research Council of Canada for performing the ^1H NMR, Paul Amor for the α-O:54 polyclonal antiserum, Christine Dodgson for construction of the Tn*phoA* fusions, and Karen Sutherland for technical assistance.

REFERENCES

1. Lindberg, A. A., and Le Minor, L. (1984) *Methods Microbiol.* **14**, 1–141
2. Reeves, P. R. (1992) *FEMS Microbiol. Lett.* **100**, 509–516
3. Keenleyside, W. J., Perry, M., MacLean, L., Poppe, C., and Whitfield, C. (1994) *Mol. Microbiol.* **11**, 437–448
4. Keenleyside, W. J., and Whitfield, C. (1995) *J. Bacteriol.* **177**, 5247–5253
5. Meier-Dieter, U., Barr, K., Starman, R., Hatch, L., and Rick, P. D. (1992) *J. Biol. Chem.* **267**, 746–753
6. Whitfield, C. (1995) *Trends Microbiol.* **3**, 178–185
7. Clarke, B. R., and Whitfield, C. (1992) *J. Bacteriol.* **174**, 4614–4621
8. Clarke, B. R., Bronner, D., Keenleyside, W. J., Severn, W. B., Richards, J. C., and Whitfield, C. (1995) *J. Bacteriol.* **177**, 5411–5418
9. Kido, N., Torgov, V. I., Sugiyama, T., Uchiya, K., Sugihara, H., Komatsu, T., Kato, N., and Jann, K. (1995) *J. Bacteriol.* **177**, 2178–2187
10. Rick, P. D., Hubbard, G. L., and Barr, K. (1994) *J. Bacteriol.* **176**, 2877–2884
11. Weisgerber, C., and Jann, K. (1982) *Eur. J. Biochem.* **127**, 165–168

12. Szabo, M., Bronner, D., and Whitfield, C. (1995) *J. Bacteriol.* **177,** 1544–1553
13. Bronner, D., Clarke, B. R., and Whitfield, C. (1994) *Mol. Microbiol.* **14,** 505–519
14. Manoil, C., and Beckwith, J. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 8129–8133
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* 2nd Ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Neuhard, J., and Thomassen, E. (1976) *J. Bacteriol.* **126,** 999–1001
17. Meier-Dieter, U., Starman, R., Barr, K., Mayer, H., and Rick, P. D. (1990) *J. Biol. Chem.* **265,** 13490–13497
18. Rao, J. K. M., and Argos, P. (1986) *Biochim. Biophys. Acta* **869,** 197–214
19. Altschul, S. F., Gish, W., Miller, W., Myers, E., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
20. Higgins, D. G., and Sharp, P. M. (1988) *Gene (Amst.)* **73,** 237–244
21. Gaboriaud, C., Bissery, B., Benchetrit, T., and Mornon, J. P. (1987) *FEBS Lett.* **224,** 149–155
22. Shuttleworth, H., Taylor, J., and Minton, N. (1986) *Nucleic Acids Res.* **14,** 8689
23. Hitchcock, P. J., and Brown, T. M. (1983) *J. Bacteriol.* **154,** 269–277
24. Clarke, A. J., Sarabia, V., Keenleyside, W., MacLachlan, P. R., and Whitfield, C. (1991) *Anal. Biochem.* **199,** 68–74
25. Tsai, G. M., and Frasch, C. E. (1982) *Anal. Biochem.* **119,** 115–119
26. Reeves, P. R. (1994) in *Bacterial Cell Wall* (Ghuysen, J.-M., and Hakenbeck, R., ed) Vol. 27, pp. 281–317, Elsevier Science B.V., New York
27. Whitfield, C., and Valvano, M. A. (1993) *Adv. Microb. Physiol.* **35,** 135–246
28. Kuhn, H.-M., Meier-Dieter, U., and Mayer, H. (1988) *FEMS Microbiol. Rev.* **54,** 195–222
29. von Heijne, G. (1992) *J. Mol. Biol.* **225,** 487–494
30. Marolda, C. L., and Valvano, M. A. (1993) *J. Bacteriol.* **175,** 148–158
31. Bugert, P., and Geider, K. (1995) *Mol. Microbiol.* **15,** 917–933
32. Becker, A., Kleickmann, M., Keller, M., Arnold, W., and A. P. (1993) *Mol. & Gen. Genet.* **241,** 367–379
33. Glucksmann, M. E., Reuber, T. L., and Walker, G. C. (1993) *J. Bacteriol.* **175,** 7033–7044
34. Reuber, T. L., and Walker, G. C. (1993) *Cell* **74,** 269–280
35. Gotschlich, E. C. (1994) *J. Exp. Med.* **180,** 2181–2190
36. Saxena, I. M., Brown, R. M., Jr., Fevre, M., Geremia, R. A., and Henrissat, B. (1995) *J. Bacteriol.* **177,** 1419–1424
37. Paulson, J. C., and Colley, K. J. (1989) *J. Biol. Chem.* **264,** 615–618
38. Holst, O., Zahringer, U., Brade, H., and Zamojski, A. (1991) *Carbohydr. Res.* **215,** 323–335
39. Jansson, P.-E., Lindberg, A. A., Lindberg, B., and Wollin, R. (1981) *Eur. J. Biochem.* **115,** 571–577
40. Klena, J. D., Ashford, R. S. I., and Schnaitman, C. A. (1992) *J. Bacteriol.* **174,** 7297–7307
41. Fält, I. C., Schweda, E. K. H., Weintraub, H., Sturm, S., Timmis, K. N., and Lindberg, A. A. (1993) *Eur. J. Biochem.* **213,** 573–581
42. Klena, J. D., and Schnaitman, C. A. (1993) *Mol. Microbiol.* **9,** 393–402
43. Kent, J. L., and Osborn, M. J. (1968) *Biochemistry* **7,** 4396–4408
44. Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., and Götz, F. (1996) *Mol. Microbiol.* **20,** 1083–1091
45. Dillard, J. P., Vandersea, M. W., and Yother, J. (1995) *J. Exp. Med.* **181,** 973–983
46. Arrecubieta, C., Garcia, E., and Lopez, R. (1995) *Gene (Amst.)* **167,** 1–7
47. DeAngelis, P. L., Papaconstantinou, J., and Weigel, P. H. (1993) *J. Biol. Chem.* **268,** 181–184
48. Dougherty, B. A., and van de Rijn, I. (1994) *J. Biol. Chem.* **269,** 169–175
49. Knirel, Y. A., and Kochetkov, N. K. (1994) *Biochem. (Moscow)* **59,** 1325–1383
50. Jann, K., Goldemann, G., Weisgerber, C., Wolf-Ullisch, C., and Kanegasaki, S. (1982) *Eur. J. Biochem.* **127,** 157–164
51. Sinnott, M. L. (1990) *Chem. Rev.* **90,** 1171–1202
52. Clarke, A. J. (1996) *Biodegradation of Cellulose: Enzymology and Biotechnology,* pp. 113–189, Technomic Publishing Company, Inc., PA.
53. Saxena, I. M., Kudlicka, K., Okuda, K., and Brown, O. M., Jr. (1994) *J. Bacteriol.* **176,** 5735–5752
54. Orlean, P., Albright, C., and Robbins, P. W. (1988) *J. Biol. Chem.* **263,** 17499–17507
55. Lidholt, K., Fjelstad, M., Jann, K., and Lindahl, U. (1994) *Carbohydr. Res.* **255,** 87–101
56. Vann, W. F., Schmidt, M. A., Jann, B., and Jann, K. (1981) *Eur. J. Biochem.* **116,** 359–364
57. Nagahashi, S., Sudoh, M., Ono, N., Sawada, R., Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Arisawa, M., and Yamada-Okabe, H. (1995) *J. Biol. Chem.* **270,** 13, 961–967
58. Liu, D., Cole, R. A., and Reeves, P. R. (1996) *J. Bacteriol.* **178,** 2102–2107
59. Saxena, I. M., Lin, F. C., and Brown, R. M., Jr. (1990) *Plant Mol. Biol.* **15,** 673–683
60. Matthysse, A. G., White, S., and Lightfoot, R. (1995) *J. Bacteriol.* **177,** 1069–1075
61. DeAngelis, P. L., Papaconstantinou, J., and Weigel, P. H. (1993) *J. Biol. Chem.* **268,** 568–571
62. Maharaj, R., May, T. B., Wang, S.-K., and Chakrabarty, A. M. (1993) *Gene (Amst.)* **136,** 267–269
63. Bulawa, C. E., Slater, M., Cabib, E., Au-Young, J., Sburlati, A., Adair, W. L. J., and Robbins, P. W. (1986) *Cell* **46,** 213–225
64. Bulawa, C. E. (1992) *Mol. Cell. Biol.* **12,** 1764–1776
65. Chen-Wu, J. L.-P., Zwicker, J., Bowen, A. R., and Robbins, P. W. (1992) *Mol. Microbiol.* **6,** 497–502
66. Smits, P. H. M., de Hann, M., Maet, C., and Grivell, L. A. (1994) *Yeast* **10,** 75–80
67. Sudoh, M., Nagahashi, S., Doi, M., Ohta, A., Takagi, M., and Arisawa, M. (1993) *Mol. & Gen. Genet.* **241,** 351–358
68. Valdivieso, M. H., Mol, P. C., Shaw, J. A., Cabib, E., and Duran, A. (1991) *J. Cell Biol.* **114,** 101–109
69. Reizer, J., Reizer, A., and Saier, M. H. (1992) *Protein Sci.* **1,** 1326–1332
70. Fernández-López, M., D'Haeze, W., Mergaert, P., Verplancke, C., Promé, J.-C., Van Montagu, M., and Holsters, M. (1996) *Mol. Microbiol.* **20,** 993–1000
71. Dénarié, J., and Cullimore, J. (1993) *Cell* **74,** 951–954
72. Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Pome, J. C., and Dénarié, J. (1990) *Nature* **344,** 781–784
73. Cabib, E., Bowers, B., and Roberts, R. L. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 3318–3321
74. Danaher, R. J., Levin, J. C., Arking, D., Burch, C. L., Sandlin, S., and Stein, D. C. (1995) *J. Bacteriol.* **177,** 7275–7279
75. Fleischmann, R. D., *et al.* (1995) *Science* **269,** 496–512
76. Göhmann, S., Manning, P. A., Alpert, C.-A., Walker, M. J., and Timmis, K. N. (1994) *Microb. Pathog.* **16,** 53–64
77. Skurnik, M., Venho, R., Toivanen, P., and al-Hendy, A. (1995) *Mol. Microbiol.* **17,** 575–594
78. Stingele, F., Neeser, J.-R., and Mollet, B. (1996) *J. Bacteriol.* **178,** 1680–1690
79. Holland, D., and Wolk, C. P. (1990) *J. Bacteriol.* **172,** 3131–3137
80. Collins-Emerson, J. M., Terzaghi, E. A., and Scott, D. B. (1990) *Nucleic Acids Res.* **18,** 6690
81. Törük, I., Kondorosi, E., Stepkowski, T., P-sfai, J., and Kondorosi, A. (1984) *Nucleic Acids Res.* **12,** 9509–9524
82. Egelhoff, T. T., Fisher, R. F., Jacobs, T. W., Mulligan, J. T., and Long, S. R. (1985) *DNA (N. Y.)* **4,** 241–248
83. Rossen, L., Johnston, A. W., and Downie, J. A. (1984) *Nucleic Acids Res.* **12,** 9497–9508
84. Vasquez, M., Davalos, A., de las Penas, A., Sanchez, F., and Quinto, C. (1991) *J. Bacteriol.* **173,** 1250–1258
85. Geremia, R. A., Mergaert, P., Geelen, D., van Montagu, M., and Holsters, M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 2669–2673
86. Goethals, K., Gao, M., Tomekpe, K., van Montagu, M., and Holsters, M. (1989) *Mol. & Gen. Genet.* **219,** 289–298
87. Rosa, F., Sargent, T. D., Rebbert, M. L., Michaels, G. S., Jamrich, M., Grunz, H., Jonas, E., Winkles, J. A., and Dawid, I. B. (1988) *Dev. Biol.* **129,** 114–123